

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on:

September 30, 2004, 10:44:58 ; Search time 7578 Seconds  
(without alignments)  
12239.632 Million cell updates/sec

Title: US-09-900-751-1

Perfect score: 3106 catggtagacggctgccccgg.....ttaaaaaaa.....aaaaaa 3106

Sequence: IDENTITY\_NUC Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \* 1: em\_estba: \* 2: em\_estbm: \* 3: em\_estin: \* 4: em\_estmu: \* 5: em\_estov: \* 6: em\_estpl: \* 7: em\_estro: \* 8: em\_htc: \* 9: gb\_est1: \* 10: gb\_est2: \* 11: gb\_hic: \* 12: gb\_est3: \* 13: gb\_est4: \* 14: gb\_est5: \* 15: em\_estfun: \* 16: em\_stom: \* 17: em\_estm: \* 18: em\_gss\_inv: \* 19: em\_gss\_pln: \* 20: em\_gss\_vrt: \* 21: em\_gss\_fun: \* 22: em\_gss\_man: \* 23: em\_gss\_mus: \* 24: em\_gss\_pro: \* 25: em\_gss\_rnd: \* 26: em\_gss\_phg: \* 27: em\_gss\_vrl: \* 28: gb\_gss1: \* 29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**RESULTS**

Result	No.	Score	Query	Length	DB	ID	Description
1	3102.8	99.9	3248	11	BC008514	BCC008514	RESULT 1
2	3087.5	99.4	4017	11	AK052738	Mus muscu	LOCUS
3	1790	57.6	2348	29	AY419858	AK052738	DEFINITION
4	1558.6	50.2	2487	29	AY419856	AY419858	ACCESSION
							VERSION
							KEYWORDS
							HTC
							SOURCE
							ORGANISM
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Etherida; Rodentia; Sciurognathi; Muridae; Murina; Mus. (bases 1 to 3248)

**REFERENCES**

1. Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F., Drablenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tohiti, Y., Carninci, P., Prange, C., Raha, S.S., Loqueland, N.A., Peters, G.J., Worley, K.C., Hale, D., Garcia, A.M., Gay, J.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

TITLE	Bouffard,G.-G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Schnarch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBLISHED	12/27/93
REFERENCE	24 (bases 1 to 3248)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cggabps1@mail.nih.gov">cggabps1@mail.nih.gov</a> Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a> Contract: (Dickson, Mark) <a href="mailto:mdcd@paxil.stanford.edu">mdcd@paxil.stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	source 1. -3248 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10990" /clone="IMage:3481914" /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue." /clone_lib="NCI CGP_Mam5" /lab_host="DB110B" /note="vector: PCMV-SPORT6"
ORIGIN	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Series: IMK Plate: 7 Row: h Column: 16 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6677898 This clone has the following problem: no 5' EST match. Location/Qualifiers
Query Match	99 %; Score 3102.8; DB 11; Length 3248;
Best Local Similarity	99 %; Pred. No. 0;
Matches	3104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DY	1 CATGGTAGCGGTGCCGGAGGACACGGCTCTAGACCGCGATGGACCGCGAA 60
Db	111 CATTGGTAGCGGTGCCGGAGGACACGGCTCTAGACCGCGATGGACCGCGAA 170
QY	61 CCATGGGTAGCACTGGGGCGCAAGCCGGAGGGCTCTAGACCTTCGGCGGGAC 120
Db	171 CCATGGGTAGCACTGGGGCGCAAGCCGGAGGGCTCTAGACCTTCGGCGGGAC 230
QY	121 TCAAGTACAATCCCGCTAGAGAACATGAATGCTTGAGGAGGGTGGAGTTCTGC 180
Db	231 TCAAGTACAATCCCGCTAGAGAACATGAATGCTTGAGGAGGGTGGAGTTCTGC 290
QY	181 CTGGGACAGTGGCAGAACATGGAGAGGAGGCCAGGGCTGGAGGCTGCTGGAG 240
Db	291 CTGGGACAGTGGCAGAACATGGAGAGGAGGCCAGGGCTGGAGGCTGCTGGAG 350
QY	241 CAGTGGTGTAGCTCTCTGCTCTCCATGGCTGGCTCTGGTGCGACTCC 300
Db	351 CAGTGGTGTAGCTCTCTGCTCTCCATGGCTGGCTCTGGTGCGACTCC 410
QY	301 ATTATGGAAATGTCGGTTCAAAGAGCTCTCATGGCATCTGGAGTACAAATGAGA 360

Db	411	ATTACGGGATGTCGGGTCAAAGATGCTCAATGCCATCTGAGGATCAAATGAA	470
QY	361	TCTTCTGGATGGTATGAGAACCTCACCGGTTATCAGGCTGAGGAGCTGGG	420
Db	471	TCTTCTGGATGGTATGAGAACCTCACCGGTTATCAGGCTGAGGAGCTGGG	530
QY	421	TGAAGGAGGGCTGAGCTGAGAATGAGTCCTGAGGAGCTGAGGAGCTGGG	480
Db	531	TGAAGGAGGGCTGAGCTGAGAATGAGTCCTGAGGAGCTGAGGAGCTGGG	590
QY	481	AGTGGGCTGTAACCTGCCTTAGTGAGGGAGGCTGAGGAGCTGGG	540
Db	591	AGTGGGCTGTAACCTGCCTTAGTGAGGGAGGCTGAGGAGCTGGG	650
QY	541	GCATCCGCCAACACTTGGCGAGAGGTTATGAGGAGCTGGG	600
Db	651	GCATCCGCCAACACTTGGCGAGAGGAGCTGGG	710
QY	601	CATGGCCACCCGAGGACGGGACTQAACTCTTCGCTAACATCTGGAGACCTTC	660
Db	711	CATGGCCACCCGAGGACGGGACTQAACTCTTCGCTAACATCTGGAGACCTTC	770
QY	661	CCATGGACCCAGAAATGCTGAGAGACTCAGGACAAAGCTGAGTTGCCATG	720
Db	771	CCATGGACCCAGAAATGCTGAGAGACTCAGGACAAAGCTGAGTTGCCATG	830
QY	721	CCCATGGAGCAGCAGCTGAGCTTAACTCTTCGCTAACATCTGGAGACCTTC	780
Db	891	CCCATGGAGCAGCAGCTGAGCTTAACTCTTCGCTAACATCTGGAGACCTTC	950
QY	831	CCCATGGAGCAGCAGCTGAGCTTAACTACCCCTGGCTCCACAGTCCTAACCG	890
QY	781	CGCATCCCGCTGGCGTGGAGGCTGAGGAGGACGGGACTCTGGTGAGCTCACCT	840
Db	901	CGCATCCCGCTGGCGTGGAGGCTGAGGAGGACGGGACTCTGGTGAGCTCACCT	900
QY	841	TCCGAAGCTTGTGAGCTCCTGGTGTAGGAGATGGAGTGAACCTGGGACCGTGTAG	900
Db	951	TCCGAAGCTTGTGAGCTCCTGGTGTAGGAGATGGAGCTGGTACCGTGAT	1010
QY	901	ATAGCTGGCCATGAAACCCACCGCTGTGGCGGCTGTGGCACCTCTCACCT	960
Db	1071	CCTACACCTGACTTCTCTCCAGAACGCTTCTGTGAGTAAACATA	1130
Db	1021	CTGACCGGGAGATCTCGCTTGAGGGACTTTCTCAGTGCCAAGATGGAGCT	1080
QY	1131	CTGACCGGGAGATCTCGCTTGAGGGACTTTCTCAGTGCCAAGATGGAGCT	1190
Db	1081	GTTGGGCTTTTGAGTGTACACCCAAGGGACATTAGAGCCCTTACTATCCAGGCACT	1140
QY	1191	GTTGGGCTTTTGAGTGTACACCCAAGGGACATTAGAGCCCTTACTATCCAGGCACT	1250
Db	1141	ACCGGCCAACATCACTGCACTGGAATTCAGGTGCAACACCGAACGTAAGG	1200
Db	1251	ACCGGCCAACATCACTGCACTGGAATTCAGGTGCAACACCGAACGTAAGG	1310
QY	1201	TGGCTCAACTCTCATCTGGGACCCAACTGAGTACCGTGGCTCTGACCAAGG	1260
Db	1311	TGGCTCAACTCTCATCTGGGACCCAACTGAGTACCGTGGCTCTGACCAAGG	1370
QY	1221	GCACACGCGAGATTAAGTCCACTTCATTGAGTACCGAACGGACCGGT	1380
QY	1261	ACTATGTTGGAGATCAACGGGAGAGTACTGGGTGAGGGCTGAGGAGCTGGG	1320
Db	1431	GCACACGCGAGATTAAGTCCACTTCATTGAGTACCGAACGGACCGGT	1490
QY	1371	ACTATGTTGGAGATCAACGGGAGAGTACTGGGTGAGGGCTGAGGAGCTGGG	1430
Db	1381	TCTTAGCTGAGTACCTCTACGACTCCACGACGGCTGAGTGTGAGTGTGCA	1440



JOURNAL	Genome Res.	10 (11), 1757-1771 (2000)
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium, Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature	409, 685-690 (2001)
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team, Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature	420, 563-573 (2002)
REFERENCE	6	(bases 1 to 4017)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanasaki,T., Hara,A., Hasizume,W., Hayasida,K., Hayatsu,N., Hiramoto,K., Hirokane,T., Hori,F., Itohaki,K., Ishii,Y., Itoh,M., Kogawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kohlma,Y., Kondo,S., Komio,H., Kouda,M., Koya,S., Kurihara,T., Miyazawa,T., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohbayo,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibaoka,K., Shinagawa,A., Shiraki,T., Takeda,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Muramatsu,M., and Hayashizaki,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-ref@gscc.rikен.go.jp, URL: http://genome.gsc.rikен.go.jp/), Tel: 1-45-503-9222, Fax: 81-45-503-9216)	
COMMENT	cdNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken. Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.	
URL	http://genome.gsc.rikен.go.jp/	
URL	http://fantom.gsc.rikен.go.jp/	
FEATURES		
Source	1.	4017 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:D630041B06" /db_xref="MGI:12422676" /db_xref="taxon:10090" /clone="D630041B06" /tissue_type="kidney" /clone_Tib="RIKEN full-length enriched mouse cDNA library" /der_stage="0 day, neonate" 174. .2741 /note="unnamed protein product; putative suppression of tumorigenicity 14 (colon carcinoma) (MGI:MGI:133881, GB NM_011176, evidence: BLASTN, 100%, match=720)" /codon_start=1 /protein_id="BAC3512B.1" /db_xref="G1:1342937" /translation="MGSNRGRKAGGGQDFGAGLKNRLENNGPFGCVELPANNA /KVEKRGPRRWWVVLVWLFSLISLMAWLFHFRTRVVRQVKFNGHLTTEFWESL DAYERSTETSLASQKKEALYKLEIYVPGYPKHRSVAFSEVILAYWSEF IPPHAEKVERAMAVRUVTPRALKSFVITSVAPPIDERMLQTQDNCSFAL
CDS		

RAHGAAMWTFITPFGKPSVYPAKRCOMVLGRGIAUSVWVQFVLTUFTDVRDCEAFCFOL  
TYUDSLSRPMPEHAWVRAGLCSVSYNLTFLSSONVFLUTTNTDWRGEAFEFOL  
PKMSSCGCFLSDQTGTFSSPYFGHPINPNTCWNKVNENRNUKVRLFVYDPNV  
PYGSCTKDVYEVNGEKCGERSOPSVVNSNSKTVHFSDHTSYTCKDQGAKYUVDN  
DPCGMMTCRBCIRKELCDWADCEYDSDYRCRDAHPTCKDQGAKYUVDN  
DASVNDGDSBEGSCGAGSFCSNGKLQLPOSQKNGDNCNGDSBASCDSTNV  
CPTKTYSCONGLISKENPEODCKTDGSDSERKNCDFRSFQKAROVTGGTMADEG  
EHEFWMDVQHLSLQDPLWYSAHCQDDFKYSDYTMHMLGGLDOS  
EHWQVSHALGCHLQASLSPDWYSAHCQDDFKYSDYTMHMLGGLDOS  
KEASAGSQEOLKURITTHESENDFTDIALLEKSESYTVEPRICLUDATHVFP  
AGKJLWITWGHTKEGGTALLQKGERVINGOTCEIDMPOQITPRMCVUGSFGV  
DSCQGDSSGPSSAEDKGMRFOAVVSMEGCAQRNPQPVYTRLPVVRDNIKETGV"

QY	841	TCCGAGCTTGTATGATGATGGCATGGACCTGACCGTGTACCGTGATG 900	Db
Db	952	TCCGAGCTTGTATGATGATGGCATGGACCTGACCGTGTACCGTGATG 1011	Db
QY	901	ATAGCGTGGCCCTTGGAACCCCGCTGGCGGTGTGGCCACCTTCACCT 960	Db
Db	1012	ATAGCGTGGCCCTTGGAACCCCGCTGGCGGTGTGGCCACCTTCACCT 1071	Db
QY	961	CCTRACAACCTGACTTCTCTCCAGAACGCTTCTGTCAAGCTGATAACATA 1020	QY
Db	1072	CCTRACAACCTGACTTCTCTCCAGAACGCTTCTGTCAAGCTGATAACATA 1131	Db
QY	1021	CTGACCGGGGACATCTCTGCCTTGGCCACCTTCTTCAGCTGCCAGATGACCT 1080	QY
Db	1132	CTGACCGGGGACATCTCTGCCTTGGCCACCTTCTTCAGCTGCCAGATGACCT 1191	Db
QY	1081	GTCGGGGCTTTAGTGTGACACCAAGGACATTAGZGCCCTACTATCCGGCCACT 1140	QY
Db	1192	GTCGGGGCTTTAGTGTGACACCAAGGACATTAGZGCCCTACTATCCGGCCACT 1251	Db
QY	1141	ACCGCCCACATCACTCACACATGGAAATCAGTGGCCAAACACGGAAACGTGAGG 1200	QY
Db	1252	ACCGCCCACATCACTCACACATGGAAATCAGTGGCCAAACACGGAAACGTGAGG 1311	Db
QY	1201	TGGCCTTAACCTCTCTATCTCTGGTGGACCCACGGTACCTGGCTCTGCCACAGG 1260	QY
Db	1312	TGGCCTTAACCTCTCTATCTCTGGTGGACCCACGGTACCTGGCTCTGCCACAGG 1371	Db
QY	1261	ACTATGTTGAGATCAACGGGAGAACATCTGTTGGACCCACAGTGGCTCTACCT 1320	QY
Db	1372	ACTATGTTGAGATCAACGGGAGAACATCTGTTGGACCCACAGTGGCTCTACCT 1431	Db
QY	1321	GCAACACGAGGAGATACAGTCCACTTCATCTGAACTCTGTCACGGACACCGGT 1380	QY
Db	1432	GCAACACGAGGAGATACAGTCCACTTCATCTGAACTCTGTCACGGACACCGGT 1491	Db
QY	1381	TCTCTAGTGTGAGTACCTCTCTACGGTACAGTCCACGGTACAGTGGCTCTACGGT 1440	QY
Db	1492	TCTCTAGTGTGAGTACCTCTCTACGGTACAGTGGCTCTACGGTACAGTGGCTCTACGGT 1551	Db
QY	1441	AGACTGACGGGATCCGAAGGAACTGCGCTGGACGGTGGAGACTCCCCGATT 1500	QY
Db	1552	AGACTGACGGGATCCGAAGGAACTGCGCTGGACGGTGGAGACTCCCCGATT 1611	Db
QY	1501	ATAGCGTGGGGTTACTGCCATGATGATGCCACCCACCGTCAAGTCAAACCGT 1560	QY
Db	1612	ATAGCGTGGGGTTACTGCCATGATGATGCCACCCACCGTCAAGTCAAACCGT 1671	Db
QY	1561	TCTGCAAGCCCTCTCTGGCTCTGAGACAGTGTGACAGACTGTGGGACCGAAGGACG 1620	QY
Db	1672	TCTGCAAGCCCTCTCTGGCTCTGAGACAGTGTGGGACCGAAGGACG 1731	Db
QY	1621	AGGAGGCTGCGTGTCTCTGGAGTTCAAGTGTCAATGGGAAGTGTCTCCCTC 1680	QY
Db	1732	AGGAGGCTGCGTGTCTCTGGAGTTCAAGTGTCAATGGGAAGTGTCTCCCTC 1791	Db
QY	1681	AGAGCGAGAAGTGTAAAGGAGAACATGGAGATGGCTGGAGCTGGCTGGTG 1740	QY
Db	1792	AGAGCGAGAAGTGTAAAGGAGAACATGGAGATGGCTGGAGCTGGCTGGTG 1851	Db
QY	1741	ACAGCGTGAATGTGCTCTCTGACACAAATAACCTACCGTGGCAAAATGGCTCTGTG 1800	QY
Db	1852	ACAGCGTGAATGTGCTCTCTGACACAAATAACCTACCGTGGCAAAATGGCTCTGTG 1911	Db
QY	1801	TGAGCGAGGCAACCTGTGAGTGGAGACGGACTGTAGGTGGTCCATGGTGGAGA 1860	QY
Db	1912	TGAGCGAGGCAACCTGTGAGTGGAGACGGACTGTAGGTGGTCCATGGTGGAGA 1971	Db
QY	1861	AAAATGTTGACTGTGGCTGGATCTTACAAACAGGTGCGCTGGTGGACCGA 1920	QY
Db	1972	AAAATGTTGACTGTGGCTGGATCTTACAAACAGGTGCGCTGGTGGAGACCGA 2031	Db
QY	1921	ATGGGGAGGAGGGAGGTGGCCCTGGCAGGTGGCCCTGGCCAGGGCAACT 1980	QY
QY	2032	ATGCGGACGAGGAGTGGCTCTGCGAGGTGGCTCCACGCCCTGGCCAGGGCACT 2091	Db
QY	1981	TGTTGGGGCTCTCTCATCTCTCTGACTGTGCGTGTGCTGCTGCTGCTGCTG 2100	QY
Db	2152	ATGACAAATTCAGTACTCAGACTACAGTGGCTGGCTGGCTGGCTGGCTGG 2211	Db
QY	2092	TGTTGGGGCTCTCTCATCTCTCTGACTGTGCGTGTGCTGCTGCTGCTG 2151	QY
QY	2041	ATGACAAATTCAGTACTCAGACTACAGTGGCTGGCTGGCTGGCTGGCTGG 2160	QY
Db	2212	ACGAGGAGGGAGTGGCTCTGGGTGAGGCTGAGCTGAGTGGCTGGAGCT 2271	Db
QY	2161	ACCTTCTCTCAATGATTGACCTTCAGCTGAGTGGCTGGCTGGAGT 2320	QY
Db	2332	CGTGGAGTACAGCACCCTGCTGCGCCCATGCTGCTGCTGCTGCTGCTG 2391	Db
QY	2281	CTCTGCAACGCCATTGGTACAGCTGCCAGCTGGGACACAAAGGGAGTACCGG 2340	QY
Db	2392	CTGCTGGCAACGCCATTGGTACAGCTGCCAGCTGGGACACAAAGGGAGTACCGG 2451	Db
QY	2341	CGCTGATCTGAGAAGGGTAGATCGGTGACGGTGGCTGGAGA 2400	QY
Db	2452	CGTGTACCTGAGAAGGGTAGATCGGTGACGGTGGCTGGAGA 2511	Db
QY	2401	TGCGCGAGTACACCCACCTCTGTCGCCCATGCTCTCTATGTCACCTGTC 2280	QY
Db	2512	TGCGCGAGTACACCCACCTCTGTCGCCCATGCTCTCTATGTCACCTGTC 2571	Db
QY	2461	CCTGCCAGGGTACTCTGGTCCCTCTGTCAGCGGAGAAGATGGGGATGTC 2520	QY
Db	2572	CTGCCAGGGTACTCTGGTCCCTCTGTCAGCGGAGAAGATGGGGATGTC 2631	Db
QY	2521	AGCTGATGTGGAGTGGCTGGTGAAGGCTGCTGAGGACACGGTGGACG 2580	QY
Db	2632	AGCTGATGTGGAGTGGCTGGTGAAGGCTGCTGAGGACACGGTGGAC 2691	Db
QY	2581	CTAGGCTCCCTGAGTGGCTGGACTGATCAAGAGGACACTGGGTATAGCGATGAC 2640	QY
Db	2692	CTAGGCTCCCTGAGTGGCTGGACTGATCAAGAGGACACTGGGTATAGCGATGAC 2751	Db
QY	2641	AGACASCGCACACACACCCACAGGACCCGACATGCCACCTGGTACAGGAG 2700	QY
Db	2752	AGACASCGCACACACACCCACAGGACCCGACATGCCACCTGGTACAGGAG 2811	Db
QY	2701	GAACACTGACCAACACCCACAGGACCCGACATGCCACCTGGTACAGGAG 2760	QY
Db	2812	GAACACTGACCAACACCCACAGGACCCGACATGCCACCTGGTACAGGAG 2871	Db
QY	2761	GCATCCTTGTAGGACTGAGTTCTCCAACATGGGACCCCTCAAGCTGGAGAGAC 2820	QY
Db	2872	GCATCCTTGTAGGACTGAGTTCTCCAACATGGGACCCCTCAAGCTGGAGAGAC 2931	Db
QY	2821	TTGGCTGTGAGGAGTGGCTGGCCAGGCTGGGGCAAGGGTTGATGCCACCTCCCTAGCC 2880	QY
Db	2932	TTGGCTGTGAGGAGTGGCTGGCCAGGCTGGGGCAAGGGTTGATGCCACCTCCCTAGCC 2991	Db
QY	2881	CTGACTGTGGAGTGGCTGGCCAGGCTGGGGCAAGGGTTGATGCCACCTCCCTAGCC 2940	QY
Db	3052	GGAGCCCTATGGGAGGGCTCAGGTCTCTTCAAGAGGCGCCAGGCTAGGA 3111	Db
QY	3001	ACCCAGAAAGACTGTGACTCTAAGGTGAATGTTGCTGTCGGGGGGT 3060	QY



QY	1584	TGTGAGACTGTCACAAGCACTGTTGGGACGGAGTGA         	1643	RESULT 4
Db	1302	TGTGAGACTGTCACAACACTGTTGGGACGGAGTGA         	1561	AY419856
QY	1644	GGGAATTCAGTGTCCAAATGGGAGTGTCTCCCTCAGGCC         	1703	LOCUS
Db	1362	GGGAATTCAGTGTCCAAATGGGAGTGTCTCCCTCAGGCC         	1703	DEFINITION
QY	1704	GACAACGTGAGATGGATGGCTCATGTA         	1763	Homosapiens STR gene, VIRTUAL TRANSCRIPT, Partial sequence,
Db	1422	GACAACGTGAGATGGATGGCTCATGTA         	1763	genomic survey sequence.
QY	1764	ACCAATACTACCGCTCCAAATGGCTCTCTGAGCAGGG         	1823	AY419856
Db	1482	ACCAATACTACCGCTCCAAATGGCTCTCTGAGCAGGG         	1823	VERSION
QY	1824	GATGGAGACGGAGCTGTAGCGATGCTCCGATGAGA         	1883	GSS
Db	1542	GATGGAGACGGAGCTGTAGCGATGAGA         	1883	SOURCE
QY	1884	TCCATTACCAACCGCTCGCGTGTTGGACGAA         	1943	ORGANISM
Db	1602	TCCATTACCAACCGCTCGCGTGTTGGACGAA         	1943	Homo sapiens
QY	1944	TGGAGGTAGCTCCACCCCTGGCCAGGC         	2003	Bukit-Jaya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db	1662	TGGAGGTAGCTCCACCCCTGGCCAGGC         	2003	Mammalia; Buttocks; Primates; Catarrhini; Hominidae; Homo;
QY	2004	CCTGACTGCTGGCTCTCGAGCTCTTCA         	2171	REFERENCE
Db	1722	CCTGACTGCTGGCTCTCGAGCTCTTCA         	2171	AUTHORS
QY	2064	GACTAACGATGTTGACGCC         	2171	Clark,A.G., Giansanti,S., Nielson,R., Thomas,P., Kejariwal,A.,
Db	1782	GACTAACGATGTTGACGCC         	2171	Ferreria,S., Nielson,R., Thomas,P., Kejariwal,A.,
QY	2124	GGGGTGGAGGCTGA         	2183	Todd,M.A., Tannenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Db	1842	GGGGTGGAGGCTGA         	2183	Adams,M.D., and Cargill,M.
QY	2184	TTCGACATGACATGCC         	1901	COMMENT
Db	1902	TTCGACATGACATGCC         	1901	JOURNAL
QY	2244	CGCCCATCTGCCTGATGCT         	2303	REVIEWED
Db	1962	CGCCCATCTGCCTGATGCT         	2303	2 (bases 1 to 2487)
QY	2304	ACAGGGTGGGCCACCAA         	2363	Clark,M.A., Giansanti,S., Nielson,R., Thomas,P., Kejariwal,A.,
Db	2022	ACAGGGTGGGCCACCAA         	2363	Ferreria,S., Nielson,R., Thomas,P., Kejariwal,A.,
QY	2364	ATCCGTGTCATACCA         	2423	Todd,M.A., Tannenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Db	2082	ATCCGTGTCATACCA         	2423	Adams,M.D., and Cargill,M.
QY	2424	ATGATGTTGTTGGT         	2483	Direct Submission
Db	2142	ATGATGTTGTTGGT         	2483	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
QY	2484	CCCTTGTCAAGCGGG         	2543	Rockville, MD 20850, USA
Db	2202	CCCTTGTCAAGCGGG         	2543	This sequence was made by sequencing genomic exons and ordering them
QY	2544	GAAGGTGGCTCAGAG         	2603	based on alignment.
Db	2262	GAAGGTGGCTCAGAG         	2603	Location/Qualifiers
QY	2604	TGGATTAAGAGCAC         	2630	1. .:2487
Db	2322	TGGATTAAGAGCAC         	2348	/organism="Homo sapiens"
QY				/mol_type="genomic DNA"
				/db_xref=taxon:9606"
				<1..>2487
				/gene="ST14"
				/locus_tag="HCN7030"
				ORIGIN
				Query Match
				Best Local Similarity 50.2%; Score 1558.6; DB 29; Length 2487;
				Matches 1774; Conservative 0; Mismatches 713; Indels 0; Gaps 0;
				ORIGIN
QY				144 AACATGATGCTTGGAGGAGCGTGA         
Db				1 ARAGTGAATGGCTTGGAGGAGCGTGA         
QY				204 GAGAAGGAGGCCAGGG         
Db				61 GAAAGCATGCCAGGG         
QY				264 CTCTCCATGCTGGCT         
Db				121 GCTCTGGGGATC         
QY				324 AAAGCTTCAATGG         
Db				181 AAGGTCTCAATGG         
QY				384 TCCACCTCCAGAG         
Db				241 TCCAATTC         
QY				444 TACAATGGCT         
Db				301 TACAGGGAGT         
QY				504 GAGGGAGT         







source	1. .2689	QY	660	-----	-----	659
/organism="Mus musculus"						
/mol_type="mRNA"						
/strain="C57BL/6J"						
/db_xref="PANTOM DB-D630022P10"						
/db_xref="NCI:242473"						
/db_xref="axon:10090"						
/clone="D630022P10"						
/tissue_type="kidney"						
/clone_Tib="RIKEN full-length enriched mouse cDNA library"						
1. .2689						
/note="suppression of tumorigenicity 14 (colon carcinoma) match=720)"						
ORIGIN						
Query Match	30.8%	Score 956.4;	DB 11;	Length 2689;		
Best Local Similarity	76.7%	Pred. No. 3.5e-17;	Mismatches 1;	Indels 428;	Gaps 3;	
Matches 1415; Conservative						
QY	1	CATGGAGACGACTGCCGAGGGACACCGCGCTGAGACCGCGTGCAGCCGCAA	60			
Db	110	CATGGAGACGACTGCCGAGGGACACCGCGCTGAGACCGCGTGCAGCCGCAA	169			
QY	61	CCATGGGTAGCATCGGGCGCAAGCGCGAGGGACTCTAGGACTTCGGGGAC	120			
Db	170	CCATGGGTAGCATCGGGCGCAAGCGCGAGGGACTCTAGGACTTCGGGGAC	229			
QY	121	TCAAGTACAATCCCGCTTAAGAACATGATGGCTTGAGCAGGGTGTAGTTCTGC	180			
Db	230	TCAAGTACAATCCCGCTTAAGAACATGATGGCTTGAGCAGGGTGTAGTTCTGC	289			
QY	181	CTGGGAAACAATCCAAAGAATGGAGAGGCCAGGGCTGGGTGGTGGAGTCTGC	240			
Db	290	CTGGGAAACAATCCAAAGAATGGAGAGGCCAGGGCTGGGTGGTGGAGTCTGC	349			
QY	241	CAGTGCCTGTTAGCTTCCTCTGCTTCCTCATGGCTGCTGGGACTTC	300			
Db	350	CAGTGCCTGTTAGCTTCCTCTGCTTCCTCATGGCTGCTGGGACTTC	409			
QY	301	ATTATGGAAATGGCGGGTCAAAGTCTCAATGCCATCTGAGGATCAAATGAGA	360			
Db	410	ATTATGGAAATGGCGGGTCAAAGTCTCAATGCCATCTGAGGATCAAATGAGA	469			
QY	361	TCTTCTGGATTCGGTATGAGAACTCCACCTCACAGAGTTATCAG-CCTGGCCAGCCAG	419			
Db	470	TCTTCTGGATTCGGTATGAGAACTCCACCTCACAGAGTTATCAGCCCTGGCCAGCAG	529			
QY	420	GTGAAGGAGGGCTGAGCTCTGTTACAATGAGCTCTGGTCTTACAGAGG	479			
Db	530	GTGAAGGAGGGCTGAGCTCTGTTACAATGAGCTCTGGTCTTACAGAGG	589			
QY	480	AAGCTGGCTGAACTGGCTTAGTGAGGGAGTGTCTACTACCTGGCAGGTC	539			
Db	590	AAGCTGGCTGAACTGGCTTAGTGAGGGAGTGTCTACTACCTGGCAGGTC	649			
QY	540	AGCATCCCCACACTGGAGAGGGTCTGGCTGAGCTGAGCTGGCTGAGCTG	659			
Db	710	AGCATCCCCACACTGGAGAGGGTCTGGCTGAGCTGAGCTGGCTGAGCTG	769			
QY	660	-	659			
Db	770	CGTGTAGTGGGGCAACCAAGGGTGGAAATGTGTGGCCAGCCTGAAAGGGGGGCT	829			
QY	660	-	659			
Db	830	GGGAGAGCCACACTCACAGGCCAGGTCTGGAGGACTCTGCTGAGGTGCACT	889			
QY	697	-	697			696
Db	1010	GGTGTACATGGGAGAAAGTGTCTGGCCCTTAAGCCATGGGTCCAAAGTCCTTC	1129			
QY	697	-	697			
Db	1110	TTCGTGTTACCTGAGTTCGGAGCTAGACGATGGGGTAGGCTAGGGGG	1189			
QY	697	-	697			
Db	1190	CGGGGCGCTGAGGGTTCACGCCTTCCTCCATCTCCACAGACAGCTGCACTTC	1249			
QY	713	CCTGCATGCCATGGCTGACGGTGAACGGCTCTACCCCTGGCTCCCAAACACTCC	772			
Db	1250	CCTGCATGCCATGGCTGACGGTGAACGGCTCTACCCCTGGCTCCCAAACACTCC	1309			
QY	773	CTACCCGGATGCCGCTGCCAGGGCTGGGGCTGCGGGGAGGCCAGCAGCTGCTGAG	832			
Db	1310	CTACCCGGATGCCGCTGCCAGGGCTGGGGCTGCGGGGAGGCCAGCAGCTGCTGAG	1369			
QY	833	CCTCACCTCCGAAAGCTTGATGTCGCTCCCTGTGATGAGCTGAGCTGCTGAG	892			
Db	1370	CCTCACCTCCGAAAGCTTGATGTCGCTCCCTGTGAGCTGAGGATGGCTGAGCTGCA	1429			
QY	893	CCTGATGATAGCCCTGAGGCCATGGCCAGCTGCTGAGCTGCTGAGCTGCTGAG	952			
Db	1430	CCTGATGATAGCCCTGAGGCCATGGCCAGCTGCTGAGCTGCTGAGCTGCTGAG	1489			
QY	953	CTCACCCCTACACCTGACTTCTCTCTCCAGAACGCTCTCCAGAACGCTCTCCAGT	1012			
Db	1490	CTCACCCCTACACCTGACTTCTCTCTCCAGAACGCTCTCCAGAACGCTCTCCAGT	1549			
QY	1013	AACCAATCTGACCGGGGACATCTGGCTTGTGAGCTTCTCCAGAACGCTCTCCAGT	1072			
Db	1550	AACCAATCTGACCGGGGACATCTGGCTTGTGAGCCACTCTCCAGAACGCTCTCCAGT	1609			
QY	1073	GAGCAGCTGCTGGGGCTTTTGAGTGACCCCAGGGACATTACGGCTACTATCC	1132			
Db	1610	GAGCAGCTGCTGGGGCTTTTGAGTGACCCCAGGGACATTACGGCTACTATCC	1669			
QY	1133	AGGCCACTACCCGCCACATCACTGAGGATATCAAGGGCTAACACCGGAA	1192			
Db	1670	AGGCCACTACCCGCCACATCACTGAGGATATCAAGGGCTAACACCGGAA	1729			
QY	1193	CCTGAGGCTGCTCAACTCTCTGAGGACCCACGTTACAGCTGCTGAGGTCCAGTTGT	1252			
Db	1730	CCTGAGGCTGCTCAACTCTCTGAGGACCCACGTTACAGCTGCTGAGGTCCAGTTGT	1789			
QY	1253	CACCAAGACTATGGAGATCAACGGGAGAAGTACTCGGTGAGGAGTCCAGTTGT	1312			
Db	1790	CACCAAGACTATGGAGATCAACGGGAGAAGTACTCGGTGAGGAGTCCAGTTGT	1849			
QY	1313	GGTGTACATGGGAGAAAGTGTCTGGCCAGCCTGAAAGGGGGCT	1416			
Db	1850	GGTGTACATGGGAGAAAGTGTCTGGCCAGCCTGAAAGGGGGCT	1909			
QY	1373	CACGGGTTCTAGTGTACACTCTCTGAGGACCTGAGCTACAGCTGAGACGAA	1953			
Db	1910	CACGGGTTCTAGTGTACACTCTCTGAGGACCTGAGACGAA	1953			

REFERENCE	BQ855511	944 bp mRNA linear EST 16-AUG-2002
LOCUS	BQ855511	AGENCourt-8748027 NIH_MGC_130 Mus musculus cDNA clone IMAGE:63335237
DEFINITION	5', mRNA sequence.	
ACCESSION	BQ855511	
VERSION	BQ855511.1	GI:22277569
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
AUTHORS	Bukay-Yoata; Metzoco; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.	
TITLE	NIH-MGC http://mgc.ncbi.nih.gov/	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabps-f@mail.nih.gov Tissue Procurement: Mark Macdonchie, Ph.D. and Nancy L. Freeman, Ph.D.	
CDNA Library Preparation: ResGen, Invitrogen Corp		
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNNL)		
DNA Sequencing by: Agencourt Bioscience Corporation		
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNNL at:		
http://image.llnl.gov		
Plate: LNM13797 row: F column: 06		
High quality sequence stop: 623.		
Location/Qualifiers		
1.. 944		
/organism="Mus musculus"		
/mol_type="mRNA"		
/db_xref="taxon:10090"		
/clone="IMAGC:6335237"		
/lab_host="NH10B (phage-resistant)"		
/clone_id="NIH_MGC_130"		
/note="Organ: otcysts; Vector: pCMV-SPORT6.1; Site: 1: EcoRV; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by RasGen, Invitrogen Corp. Note: this is a NIH_MGC Library."		
ORIGIN		
Query Match		
Best Local Similarity	27.2%	Score 843.8; DB 13; Length 944;
Matches	885;	Conservative 0; Mismatches 32; Indels 2; Gaps 2;
QY		
30	GCGGTGAGCCGGGATCGAACCCCAAACCATGGTAGCAATGGGGACGCCAGGCC 89	
Db	1 GCGTGTGAGCCGGGATCGAACCCCAAACCATGGTAGCAATGGGGACGCCAGGCC 60	
QY		
90	GGAGGGGGCTCTAGAACTCGGGGGACTCAAGTACAACCTCCGGCTAGAACATG 149	
Db	61 GGAGGGGGCTCTAGAACTCGGGGGACTCAAGTACAACCTCCGGCTAGAACATG 120	
QY		
150	AATGGCTTGAGGAAGGGTGGAGTTCTGGCTGGAGACATGCCAAGTGGAGAG 209	
Db	121 ATGGCTTGAGGAAGGGTGGAGTTCTGGCTGGAGACATGCCAAGTGGAGAG 180	
QY		
210	CGAGGCCCGAGGGCTGGGGCTGGCTGGAGCTGGCTGGCTGGCTGGCTGGCTGG 269	
Db	181 CGAGGCCCGAGGGCTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 240	
QY		
270	CTCATGGCTGAGTCAGGATCACAAATGAGACTTCTGGATGGTGAAGACTCCACC 329	
Db	241 CTCAATGGCTGAGTCAGGATCACAAATGAGACTTCTGGATGGTGAAGACTCCACC 300	
QY		
330	TTCAATGGCCACCTGGAGGATCACAAATGAGACTTCTGGATGGTGAAGACTCCACC 389	
301	TTCAATGGCCACCTGGAGGATCACAAATGAGACTTCTGGATGGTGAAGACTCCACC 360	
QY		
390	TCCAAAGAGTTATCAGCTGGACCGAGCTGGACTTCCATTATGGAAATGGCGGTCAAAAGTC 449	
Db	361 TCCAAAGAGTTATCAGCTGGACCGAGCTGGACTTCCATTATGGAAATGGCGGTCAAAAGTC 420	
FEATURES		
source		
RESULT	8	
BI68367		
LOCUS	BI68367	884 bp mRNA linear EST 18-SEP-2001
DEFINITION	603316692FL NCI_CGAP_Mam6	Mus musculus cDNA clone IMAGE:53564875,
ACCESSION	BI68367	mRNA sequence.
VERSION	BI68367.1	GI:15650996
KEYWORDS		
SOURCE		
ORGANISM	Mus musculus (house mouse)	
Bukay-Yoata; Metzoco; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.		
REFERENCE		
Authors		
Title		
Journal		
Comment		
Email: cgabps-f@mail.nih.gov		
Tissue Procurement: Jeffrey Green M.D.		
CDNA Library preparation: Life Technologies, Inc.		
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNNL)		
DNA Sequencing by: Incyte Genomics, Inc.		
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNNL at:		
http://image.llnl.gov		
Plate: LNM11906 row: h column: 24		
High quality sequence stop: 331.		
Location/Qualifiers		
1.. 884		
/organism="Mus musculus"		
/mol_type="mRNA"		
/strain="FVB/N"		
/db_xref="taxon:10090"		
/clone="IMAGC:5356487"		
/sex="female, virgin"		

/tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"

/clone\_lib="NCI\_CGAP Mam6"

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Jeffrey Green, M.D., NIH

## ORIGIN

Query Match 26.8%; Score 833.6; DB 12; Length 884;  
 Best Local Similarity 98.6%; Pred. No. 4.8e-153;  
 Matches 872; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Oy 1165 GGAATATCAGGGGCCAACACCGAACAGCTGAGGTCGCCTCAACTCTCTATCTGG 1224  
 Db 1 GGAATATCAGGTGCCAACACCCGAACTGAGGTCGCCTCAACTCTCTATCTGG 60  
 Qy 1225 TGGACCCCAAAGTACCATGCTGGCTCTGCCAGGATATGAGGATACGGGAGA 1284  
 Db 61 TGGACCCCAAAGTACCATGCTGGCTCTGCCAGGATATGAGGATACGGGAGA 120  
 Qy 1285 AGTACTGGGGTGGAGGTGCCAGTTGGTGAGCACACGGTCTACTGAGTACTCTCTACG 1344  
 Db 121 AGTACTGGGTGGAGGTGCCAGTTGGTGAGCACACGGTCTACTGAGTACTCTCTACG 180  
 Qy 1345 ACTTCATCTTCATCTGTGACCGAACGGTCTACTGAGTACTCTCTACG 1404  
 Db 181 ACTTCATCTTCATCTGTGACCGAACGGTCTACTGAGTACTCTCTACG 240  
 Qy 1405 ACTCCAAAGGCGGTGCCAGGATGTCTGAGAACGAGCTGGCATCGAAAG 1464  
 Db 241 ACTCCAAAGGCGGTGCCAAAGGATGTCTGAGAACGAGCTGGCATCGAAAG 300  
 Qy 1465 AACTGCGTGTGCAAGGCTGGAGACTGCCCCGATTAGTAGTGATGAGCTACTSGCAT 1524  
 Db 301 AACTGCGTGTGCAAGGCTGGAGACTGCCCCGATTAGTAGTGATGAGCTACTSGCAT 360  
 Qy 1525 GCAATGCCACCAACCAGTCAGTGCAAAACCGAGTCTGGAAAGCCCTCTCTGGCT 1584  
 Db 361 GCAATGCCACCAACCAGTCAGTGCAAAACCGAGTCTGGAAAGCCCTCTCTGGCT 420  
 Qy 1585 GTGACAGTGTAAAGTGTGGGACCGAAAGCAGGGAGGGTGGAGCTGGCT 1644  
 Db 421 GTGACAGTGTAAAGTGTGGGACCGAAAGCAGGGAGGGTGGAGCTGGCT 480  
 Qy 1645 GAGTTCAAGTGTCAATGGAAGTGTCTCTCTAGAGCAGAAGTGTATGAAAG 1704  
 Db 481 GGAGTTCAAGTGTCTCTCTAGAGCAGAAGTGTATGAAAG 540  
 Qy 1705 ACACACTTGAGATGGCTCTGAGGCTTATGTCAGAGGGTAATGTCCTCTCTCTGCA 1764  
 Db 541 AACACTGTGGAGATGGCTCTGAGGCTTATGTCAGAGGGTAATGTCCTCTCTGCA 600  
 Qy 1765 CAAATACTACCTACCGTCCAAATGGCTCTCTGAG 1823  
 Db 601 CCAAATATACCTACCGTCCAAATGGCTCTCTGAG 660  
 Qy 1824 GATGGGAGAGGGACTGTGGATGCTCCATGAGAACACTGTCAGGGCTGCA 1883  
 Db 661 GATGGGAGAGGGACTGTGGATGCTCCATGAGAACACTGTCAGGGCTGCA 720  
 Qy 1884 TCTTCTTACCAACAGGCTCGG-TGTTGGGAGGAATGGGGAGGGACTGTGGC 1942  
 Db 721 TCTTCTTACCAACAGGCTCGG-TGTTGGGAGGAATGGGGAGGGACTGTGGC 780  
 Qy 1943 CTGGAGGTGTGGCTCACGCGCTGGGAGGGACTGTGGCTGCA 2045  
 Db 781 CTGGAGGTGTGGCTCACGCGCTGGGAGGAATGGGGAGGGACTGTGGC 840  
 Qy 2002 CTCTGACTGCTGGCTCGAGCTCATCTTGTGAGGATGAC 2045  
 Db 841 CTCTGACTGCTGGCTCGAGCTCATCTTGTGAGGATGAC 884

## RESULT 9

BI690615

803 bp mRNA

linear

EST 18 SEP 2001

mRNA sequence.

BI690615.1

GI:15653244

## VERSION

B1690615

## KEYWORDS

Mus musculus

## ORGANISM

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

## REFERENCE

1 (bases 1 to 803)

## AUTHORS

NIH-MGC

<http://mgc.ncbi.nlm.nih.gov/>

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-f@mail.nih.gov](mailto:cgapbs-f@mail.nih.gov)

Tissue Procurement: Jeffrey Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LILN at:

<http://image.llnl.gov>

Plate: LILN1900 row: a column: 04

High quality sequence stop: 755.

## FEATURES

Location/Qualifiers

1. . 803

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="PRVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:5354067"

/sex="female; virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Mam6"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

## ORIGIN

Query Match 25.4%; Score 788.8; DB 12; Length 803;  
 Best Local Similarity 99.6%; Pred. No. 3e-144; 2; Indels 1; Gaps 1;  
 Matches 801; Conservative 0; Mismatches 2;

Qy 654 GcTCTCCCATGACCCAGATGCTGAGGACATCGAGAACAGCTGAGTTGCC 713  
 Db 1 GCTCTCCCATGACCCAGATGCTGAGGACATCGAGAACAGCTGAGTTGCC 60  
 Qy 714 CTGCTCCCATGTCAGGCTGACAGCTTCACTCCCTGGCTCCACAGTCCC 773  
 Db 61 CTGCTCCCATGTCAGGCTGACAGCTGACAGCTGCTCACTCCCTGGCTCCACAGTCCC 120  
 Qy 774 TACCCGGCATGCCCTGCGCAGTGGCTGCGGGAGGGAGCCGACTGTGCTGAGC 833  
 Db 121 TACCCGGCATGCCCTGCGCAGTGGCTGCGGGAGGGAGCCGACTGTGCTGAGC 180  
 Qy 834 CTCACCTCCGAGCTTGTGCTGCTGAGGAGCTGAGCTGCTGAGGAGCTGCTGAGC 893  
 Db 181 CTCACCTCCGAGCTTGTGCTGCTGAGGAGCTGAGCTGCTGAGGAGCTGCTGAGC 240  
 Qy 894 GTGTAATGAGCTGAGGCCATGAAACCCACGCTGCTGAGGAGCTGCTGAGCTGCTGAGC 953  
 Db 241 GTGTAATGAGCTGAGGCCATGAAACCCACGCTGCTGAGGAGCTGCTGAGCTGCTGAGC 300  
 Qy 954 TCACCTCTAACACCTGACTTCTCCAGAGACTCTCTCTGAGGAGCTGCTGAGCTGCTGAGC 1013

Db 301 TCACCCCTCTAACACCTGACTTCTCCCTCCAGACGCTTCAGGTGTCACCGTGATA 360

QY 1014 ACCAATACTGACGGGCACATCTGGCTTGAGCCTTCTCCAGCTGCCAGATG 1073

Db 361 ACCAATACTGACGGGCACATCTGGCTTGAGCCTTCTCCAGCTGCCAGATG 420

QY 1074 AGCACGTTGGGGCATCTGGCTTGAGTGCACCTTCTCCAGCTGCCAGATG 1133

Db 421 AGCACGTTGGGGCATCTGGCTTGAGTGCACCTTCTCCAGCTGCCAGATG 480

QY 1134 GGCCTACTACCCGCCAACATCACTCACTGACATGAAATCAAGGTGCCAACACCGAAC 1193

Db 481 GGCCTACTACCCGCCAACATCACTCACTGACATGAAATCAAGGTGCCAACACCGAAC 540

QY 1194 GTCAGAGTGGCTCAACTCTCTATCTGGAGGACCCACGCTACAGTGCTGCTGC 1253

Db 541 GTGAAGSTGCSCTCAACTCTCTATCTGGAGGACCCACGCTACAGTGCTGCTGC 600

QY 1254 ACCAAGGACTATGTTGAGATCAACGGAGAGACTCGGGTAGGAGTCCAGTTGTG 1313

Db 601 ACCAAGGACTATGTTGAGATCAACGGAGAGACTCGGGTAGGAGTCCAGTTGTG 660

QY 1314 GTCAGAGCAGCAACGGCAAGATACAGTCCACTTCATTCTGATCACTCTGACACGAC 1373

Db 661 GTCAGAGCAGCAACGGCAAGATACAGTCCACTTCATTCTGATCACTCTGACACGAC 720

QY 1374 ACCGGGTTCTCTGAGTACCTCTCTAGCTGACTCCACGACCCTGGCCAGGATGTC 1433

Db 721 ACCGGGTTCTCTGAGTACCTCTCTAGCTGACTCCACGACCCTGGCCAGGATGTC 780

QY 1434 ATCTGCAAGACTGACGGTGATC 1457

Db 781 ATCTGCAAGA-TGGACGCGTGCATC 803

RESULT 10

BQ951529 LOCUS BQ951529 DEFINITION AGENCOURT 8952916 NCI CGAP\_Co24 Mus musculus mRNA sequence. IMAGE:6475074 5', mRNA sequence.

ACCESSION BQ951529 VERSION BQ951529.1 GI:22367007

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Bukar-Yota; Metzger; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butirota; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgsabs@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at:

<http://image.liln.gov>  
Plate: LILM14013 row: d column: 19  
High quality sequence stop: 635.  
Location/Qualifiers

FEATURES source

1. -990 /organism="Mus musculus"  
/strain="FVB/N"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6475074"  
/lab\_host="DH11B (T1 phage-resistant)"  
/clone\_lib="NCI CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: oligo dt.

Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP library."

#### ORIGIN

Query Match 25.1%; Score 778.2; DB 13; Length 990; Best Local Similarity 92.9%; Pred. No. 3.5e-142; Mismatches 60; Indels 11; Gaps 10; Matches 924; Conservative 0; Mi smatches 60; Index 10;

Db 981 TTCTCCAGAACGCTCTCTGTCAGCTGATACCAATCTGACCGCCGACATCTGGC 1040

QY 1 1CCTCCAGAACGCTCTCTGTCAGCTGATACCAATCTGACCGCCGACATCTGGC 60

Db 1041 TTGAGGCCATTCTCCAGCTGCCAGATGAGCTGAGCTGAGCTGAGCTTGTGAC 1100

Db 61 TTGAGGCCATTCTCCAGCTGCCAGATGAGCTGAGCTGAGCTGAGCTTGTGAC 120

QY 1101 ACCAAGGGACATTTAGGACGCCACTATCCAGGCAACTACCGCCACATCACTGC 1160

Db 121 ACCAAGGGACATTTAGGACGCCACTATCCAGGCAACTACCGCCACATCACTGC 180

QY 1161 ACATGGATATCAAGGTTGCCACAAACCGGACAGCTGAGGAGGAGCTCAACTCTTA 1220

Db 181 ACATGGATATCAAGGTTGCCACAAACCGGACAGCTGAGGAGGAGCTCAACTCTTA 240

QY 1221 CTTGGACCCATTCTCCAGCTGCCAGATGAGCTGAGCTGAGCTGAGCTTGTGAC 180

Db 241 CTGGAGGCCATTCTCCAGCTGCCAGATGAGCTGAGCTGAGCTGAGCTTGTGAC 300

QY 1281 GAGAGTACTGGGTAGAGGCTTGTGAGGAGGAGCTGAGGAGCTGAGCTGAGCTTGTGAC 1340

Db 301 GAGAGTACTGGGTAGAGGCTTGTGAGGAGCTGAGGAGCTGAGCTGAGCTTGTGAC 360

QY 1341 GTCCACTCTCATCTGATCACTCTGAGGAGCTGAGGAGCTGAGCTGAGCTTGTGAC 1400

Db 361 GTCCACTCTCATCTGATCACTCTGAGGAGCTGAGGAGCTGAGCTTGTGAC 420

QY 1401 TAGGACTCCAAAGACCCGTCGCCAGGAGATGTCAGTGGCAAGCTGGTCATCGA 1460

Db 421 TAGACTCCAAAGACCCGTCGCCAGGAGATGTCAGTGGCAAGCTGGTCATCGA 480

QY 1461 AAGGAACCTGGGTGCGAGCTGGCTGGAGACTGCCGGAATATAGTGTGAGGTTACTGC 1520

Db 481 AAGGAACCTGGGTGCGAGCTGGCTGGAGACTGCCGGAATATAGTGTGAGGTTACTGC 540

QY 1521 CGATGCAATGCCACCAACGAGTCACGGTCACTGAGCTGAGGAGCTGAGCTG 600

Db 541 CGATGCAATGCCACCAACGAGTCACGGTCACTGAGCTGAGGAGCTGAGCTG 600

QY 1580 GGCTGTGAGCTGGCTGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTGAGCTG 600

Db 601 GGCTGTGAGCTGGCTGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTGAGCTG 660

QY 1639 CTGGTGGAGTTCAAGTGTCCATGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 780

Db 651 CTGGTGGAGTTCAAGTGTCCATGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 720

QY 1698 GGGAGGAACTCT-GGAGATGGCTGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 1754

Db 721 GGGAGGAACTCT-GGAGATGGCTGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 780

QY 1755 GTCTCTGACCAAATAACTACCTGGCTGACAAATGGCTCTCTGAGGAGCTGAGCTG 1814

Db 781 GTCTCTGACCAAATAACTACCTGGCTGACAAATGGCTCTCTGAGGAGCTGAGCTG 837

QY 1815 CCTGAGTGGTGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 1874

Db 838 CTGAGTGGTGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 895

QY 1875 GGGTGGAGCTTACCAACGGCTGGCTGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 1934

Db 895 GGGTGGAGCTTACCAACGGCTGGCTGGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTG 955

QY 1935 GAGTGGCTGGAGGAGCTGAGGAGCTGAGCTGAGGAGCTGAGCTGAGCTGAGCTG 1969

Db	956	CAGGGCCCCGTGGGGGACCTCCCGGACTCCCGGAGG 990
RESULT 11		
LOCUS	BG919617	762 bp mRNA linear EST 05-JUN-2001
DEFINITION	602823173F1 NCI_CGAP_Mam6	Mus musculus cDNA clone IMAGE:4951962 5'
REFERENCE	NIH_MGC http://mgc.nci.nih.gov/	mRNA Sequence.
AUTHORS	1 (bases 1 to 762)	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov	
FEATURES	Tissue Procurement: Jeffrey Green M.D., DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL), DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov/plate: LIAM01909 row: a column: 19 High quality sequence stop: 727.	
source	<p>1. .762</p> <p>Location/Qualifiers</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="FVB/N"</p> <p>/db_xref="taxon:0090"</p> <p>/clone="IMAGE:4951962"</p> <p>/sex="female, virgin"</p> <p>/tissue_type="infiltrating ductal carcinoma"</p> <p>/dev_stage="5 months"</p> <p>/lab_host="DH10B"</p> <p>/clone_id="NCI CGAP Mam6"</p> <p>/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sali; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"</p>	
ORIGIN		
Query Match	24.5%	Score 760.4; DB 12; Length 762;
Best Local Similarity	99.9%	Pred: No. 1.1e-138; Matches 761; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	682	AGAGGATCAGGACAACTCGCATTTGCCATGGTCAGCTGAC 741
QY	1	AGAGGATCAGGACAACTCGCATTTGCCATGGTCAGCTGAC 60
QY	742	GCTTCACTACCCCTGGCTTCCCACAGTCCTAACCGCCGATGCCGCTGCACTGG 801
QY	61	GCTTCACTACCCCTGGCTTCCCACAGTCCTAACCGCCGATGCCGCTGCACTGG 120
QY	802	TCCCTGGGGGACGCCGACTCTGGCTACCTCCGAAGCTTGATGTCCTC 861
QY	121	TCCCTGGGGGAGGCCGACTCTGGCTACCTCCGAAGCTTGATGTCCTC 180
QY	862	CCTCTGATGAGCATGGAGTGACCTGACCTGGCTGACCTGGCTGAC 921
QY	181	CCCTCTGATGAGCATGGAGTGACCTGACCTGGCTGACCTGGCTGAC 240
QY	922	CCCAAGCTGGCTGGGGGACCTCTACCCCTCTACACCTGACTTCCCT 981
QY	241	CCCAAGCTGGCTGGGGGACCTCTACCCCTCTACACCTGACTTCCCT 300
Db	982	CCTCCCGAAGCTCTCTCTGATACCAACTCGACCCGGACATCCCTGGCT 1041
Db	301	CCCTCCAGAACCTCTCTGTCACSGCAATCTGACCTCC 360
QY	1042	TTCAGGCACTTCTTCAGTGGCCAGATGAGCTGAGCGCTGCGGCGCTTTTAGTGAC 1101
Db	361	TTAGGACACTTCTTCAGTGGCCAGATGAGCTGAGCGCTGCGGCGCTTTTAGTGAC 420
QY	1102	CCCAAGGACATTAGAGCCCTATATCCAGGCCACTACCGCCAACTCAATGCA 1161
Db	421	CCAAGGACATTAGAGCCCTACTATCCAGGCCACTACCGCCAACTCAATGCA 480
QY	1162	CATGGATATCAAGGTCGCCAACACCGAACGTAAGGGGCCCTAACCTCTATC 1221
Db	481	CTGGGATATCAAGGTCGCCAACACCGAACGTAAGGGGCCCTAACCTCTATC 540
QY	1222	TGGTGGACCCAAAGGACAGCTGGGGCTCTGCCACAGGACTATGGAGATTCAAACGGG 1281
Db	541	TGGTGGACCCAAAGGACAGCTGGGGCTCTGCCACAGGACTATGGAGATTCAAACGGG 600
QY	1282	AGAAGTACTGGTGAAGAGTCCTGAGACTATGGAGATTCAAACGGG 1341
Db	601	AGAAGTACTGGTGAAGAGTCCTGAGACTATGGAGATTCAAACGGG 660
QY	1342	TCCACTTCCATTCTGATCACTGTACCGACACCGGGTCTACCTGACTACCTCTCT 1401
Db	661	TCCACTTCCATTCTGATCACTGTACCGACACCGGGTCTACCTGACTACCTCTCT 720
QY	1402	ACGACTCCACGACCGGTGCGCACGGATGTCATGTGCGAGA 1443
Db	721	ACGACTCCACGACCGGTGCGCACGGATGTCATGTGCGAGA 762

directionally, priming method: Oligo-dT. cDNA enrichment >1k bp, Average insert size 1.8K bp. Priming sequence: 5'GACTTGTCTAGATGCCGAGGCCGTC(R) 3'. Tissue contributed by, David Rowe. Library constructed by Resgen, Invitrogen Corp."

directionally, priming method: oligo-dT; cDNA enrichment: >1k bp, Average insert size 1.8k bp. Priming sequence: 5'GACTAGTCTAGTCGACGCCGCCT (T) 3'. Tissue contributed by, David Rowe. Library constructed by ResGen, Invitrogen Corp."

SOURCE MUS musculus (house mouse)  
 ORGANISM Mus musculus  
 EST. B1664081.1 GI:15578314  
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammal; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 798)  
 AUTHORS NIH-MGC <http://mcc.ncbi.nlm.nih.gov/>  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1993)  
 Contact: Robert Straussberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA library Arrayed by: The I.M.A.G.E. Consortium (LBNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLL at:  
<http://image.llnl.gov>  
 Plate: LLM1:1820 Row: 1 Column: 22  
 High quality sequence stop: 771.

QY	1567	AGCCCTCTCTGGGCTGTACAGGTCAACGAGGCTGGGAACGGAACTGACAGGGAG	1626
Db	181	AGCCCTCTCTGGGCTGTACAGGTCAACGAGGCTGGGAACGGAACTGACAGGGAG	240
QY	1627	GCTGAGCTGTCCTGGAGTTCAAGTGTCAATTGGGAAGTGTCTCCTCGAGGCC	1686
Db	241	GCTGAGCTGTCCTGGAGTTCAAGTGTCAATTGGGAAGTGTCTCCTCGAGGCC	300
QY	1687	AGAAGTGTAATGGGAAAGAACATTGGGATGGCTACAGGGCTCATGACAGCG	1746
Db	301	AGAAGTGTAATGGGAAAGAACATTGGGATGGCTACAGGGCTCATGACAGCG	360
QY	1747	TGATGTCCTCTGGCACAAATTACCTACCGTGCCTAAATGGCCCTGTGAGCA	1806
Db	361	TGATGTCCTCTGGCACAAATTACCTACCGTGCCTAAATGGCCCTGTGAGCA	420
QY	1807	AGGGCACCTGTAGTGTATGGAGAACGACTGTAGCAATGGCTCTGTGAGCA	1866
Db	421	AGGGCACCTGTAGTGTATGGAGAACGACTGTAGCAATGGCTCTGTGAGCA	480
QY	1867	GTGAGCTGAGCTGGATCTTAACTAAAGGCTGGCTGGTGTGGAGGATGCC	1926
Db	481	GTGAGCTGAGCTGGATCTTAACTAAAGGCTGGCTGGTGTGGAGGATGCC	540
QY	1927	ACGAGGGCAGTGGCCTGGAGCTGGCAGGGACTGTGGTGTGGTGTGG	1986
Db	541	ACGAGGGCAGTGGCCTGGAGCTGGCAGGGACTGTGGTGTGGTGTGG	600
QY	1987	GGCCCTGGCTCATCTCTCCCTGACTGCTGGCTCTCAGTCATGGATGACA	2046
Db	601	GGCCCTGGCTCATCTCTCCCTGACTGCTGGCTCTCAGTCATGGATGACA	660
QY	2047	AAATTTCAAGTACTAGACATCACATGAGGACCCCTCTGGACAGGAGA	2104
Db	661	AAATTTCAAGTACTAGACATCACATGAGGACCCCTCTGGACAGGAGA	719
QY	2107	GCAAAGCAGTGGCTCTGGGTGGAGCTGAACTCAAGTATCAT-CACCCACCT	2165
Db	720	GCAAAGCAGTGGCTCTGGGTGGAGCTGAACTCAAGTATCAT-CACCCACCT	779
QY	2166	TCCTCATGATT 2179	
Db	780	TCCTCATGATT 793	

421	AGCTTGACGACATGAGTCCTGTCGGTACACAGAGTGGCTGAACTG	480	COMMENT
Db			Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22, Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a>
Qy			Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hiroane,T., Hori,F., Imotani,K., Ishii,Y., Itou,M., Kawada,I., Kawai,J., Kohjima,Y., Konno,S., Kono,H., Koya,S., Miyakawa,A., Murakami,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Wakiki,K., Watanuki,A., Muramatsu,M. and Hayashiizaki,Y.
Qy			Direct Submission
Db			Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 67-67 (2001)
Qy	TGGCAGAGGACTTACAGCAACAGCTGGCTGAGTTGTCATGCCCGAG 615	600	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Qy	TGGCTGAGGACTTACAGCAACAGCTGGCTGAGTTGTCATGCCCGAG 675	720	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline (with 2000)
Qy	TGGCTGAGGACTTACAGCAACAGCTGGCTGAGTTGTCATGCCCGAG 734	779	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Db	TGGCTGAGGACTTACAGCAACAGCTGGCTGAGTTGTCATGCCCGAG 793	795	Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for further details.
Db	790 GCAGTGGGTCTGGG 795		Contact: Yoshihide Hayashizaki
RESULT 14			Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22, Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a>
BY756896			Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hiroane,T., Hori,F., Imotani,K., Ishii,Y., Itou,M., Kawada,I., Kawai,J., Kohjima,Y., Konno,S., Kono,H., Koya,S., Miyakawa,A., Murakami,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Wakiki,K., Watanuki,A., Muramatsu,M. and Hayashiizaki,Y.
LOCUS	BY756896	746 bp	mRNA linear EST 17-DEC-2002
DEFINITION	BY756896	RIKEN full-length enriched, blastocyst <i>Mus musculus</i> cDNA clone IIC003F06 3', mRNA sequence.	
ACCESSION	BY756896.1	GI:27190109	Direct Submission
KEYWORDS	EST.		Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 67-67 (2001)
SOURCE	Mus musculus (house mouse)		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
ORGANISM	Mus musculus		RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline (with 2000)
REFERENCE	I (bases 1 to 746)		Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
AUTHORS	Ozaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamamoto,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nagami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Katapin,A., Matsuda,H., Batzelov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Garibaldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Guatamacchi,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasaka,I., Kedzierski,R.M., Kling,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltsev,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Nunata,K., Okido,T., Pavan,W.J., Pertea,G., Pessele,G., Petrovsky,N., Pilati,R., Pontius,J.U., Qi,B., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sutcliffe,R., Takemoto,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirano,K., Kishikawa,T., Kono,H., Nakamura,M., Sakazume,N., Satoh,K., Shiraishi,T., Wakai,K., Aizawa,K., Arcakawa,T., Fukuda,S., Hara,A., Hashimoto,W., Imotani,K., Ishii,Y., Ichio,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasai,D., Shibusawa,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.	Genome Res. 10 (10), 1617-1630 (2000)	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 67-67 (2001)
JOURNAL	Nature 420, 563-573 (2002)		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	22354683		RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline (with 2000)
PUBMED	12466851		Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
FEATURES	source		Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for further details.
ORIGIN			Division of Experimental Animal Research in Riken contributed to please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for further details.
Query Match			Location/Qualifiers
Best Local Similarity	23.8%	Score 739.2; DB 13; Length 746;	1. . . 746
Matches	99.3%	DB 13; Length 746; Best Local Similarity 23.8%; Score 739.2; DB 13; Length 746; Matches 741; Conservative 0; Miatches 5; Indels 0; Gaps 0;	/organism="MUS musculus"
			/mol_type="mRNA"
			/strain="CS7BL/6J"
			/db_xref="taxon:10090"
			/clone="IIC003F06"
			/dev_stage="blastocyst"
			/clone_lib="RIKEN full-length enriched, blastocyst"
Qy	CTGATCCCTGCAGAGGTGAGTCGGCTCAACCAAGACACCTGTGAGGACCTCATG 2402	2343	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Db	1 CTGATCCCTGCAGAGGTGAGTCGGCTCAACCAAGACACCTGTGAGGACCTCATG 60	61	Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for further details.
Qy	CCGCAGGAGATCACCCACGATGATGTTGGTTCTCACTGGCTGAGCTGGCTGAGCTCC 2462	2403	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Db	CCGCAGGAGATCACCCACGATGATGTTGGTTCTCACTGGCTGAGCTGGCTGAGCTCC 120	61	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline (with 2000)
Qy	TGCCAAGGTGACTCTGGTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTCC 2522	2463	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Db	TGCCAAGGTGACTCTGGTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTCC 180	121	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline (with 2000)
Qy	GCTGGTGTGGTAGCTGGGTGAAGGCTGCGCTCAGAGAACGGACAGGGTGACACA 2582	2523	Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 67-67 (2001)
Db	GCTGGTGTGGTAGCTGGGTGAAGGCTGCGCTCAGAGAACGGACAGGGTGACACA 240	181	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Qy	AGCTCCTGAGTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTCC 2642	2583	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline (with 2000)
Db	AGCTCCTGAGTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTCC 300	241	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Qy	ACAGCGACCAACACCGACAGGATGCCGACATGCCGACATGCCGACATGCCGACATGCCGACAGGAGGA 2702	2643	Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to please visit our web site ( <a href="http://genome.gsc.riken.go.jp">http://genome.gsc.riken.go.jp</a> ) for further details.
Db	ACAGCGACCAACACCGACAGGATGCCGACATGCCGACATGCCGACATGCCGACATGCCGACAGGAGGA 360	301	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

TGTACCAATTCTGAACTGGGAGCGGGCCGCTCTGTTTTTTTTTTT 3'.

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fátima Bonaldo. "

Db	QY	2.421 ATCCTTAGGACTCGAGGTCTTCGAAAGGGACCCCTCAAGATGTTGAGAGAACHT 480	23.8%; Score 739.2; DB 12; Length 846;
Db	QY	2.823 GCGTCTAACGGGCCAACGCTGGGGCAAGGTTGATGGCAGCCTCCCTCTAGCCCT 2882	Best Local Similarity 97.1%; Pred No. 1.6e-134; Mismatches 18; Indels 6; Gaps 5;
Db	QY	4.81 GCGTCTAACGGGCCAACGCTGGGGCAAGGTTGATGGCAGCCTCCCTCTAGCCCT 540	Matches 816; Conservative 0; Mismatches 18; Indels 6; Gaps 5;
Db	QY	2.883 GAGCTGGTGAAGTGATGCTGTCGTCGAGAGCTCTCCAACTCTCATGAGCCCCCG 2942	1687 AGAAAGTGAATGGAGGACACTCTGGAGATGGGGCTGAGAGAGGTTCTATGTGACAGCG 1746
Db	QY	5.41 GAGCTGGTGAAGTGATGCTGTCGTCGAGAGCTCTCCAACTCTCATGAGCCCCCG 600	1 AGATCGCACAGTAGAGGACACTCTGGAGATGGGGCTGAGAGGTTCTATGTGACAGCG 60
Db	QY	2.943 GACCCCTATGGAGGAGGGCTCAGGTCACTCTTCAGGAAGGCCAGCCCTAGGAAC 3002	1747 TGAATGCGCTCTGACCAAATAACCCGGTTGCAAAATGGCCCTCTGTCAGACA 1806
Db	QY	6.01 GAGCCCTATGGAGGAGGGCTCAGGTCACTCTTCAGGAAGGCCAGCCCTAGGAAC 660	61 TGAATGCGCTCTGACCAAATAACCCGGTTGCAAAATGGCCCTCTGTCAGACA 120
Db	QY	3.003 CCCGAAAGAGTGTACTAACGGTGAATGTTTGTGTTGCAAGGGTGGTATT 3062	1807 AGGGACACCTTGAGTGTGATGGAAAGACGGCTGAGCAATGCTCGAGATAGAAAAT 1866
Db	QY	6.61 CCCAGAAAAGAGTGTACTAACGGTGAATGTTTGTGTTGCAAGGGTGGTATT 720	121 AGGGACACCTTGAGTGTGATGGAAAGACGGCTGAGCAATGCTCGAGATAGAAAAT 180
Db	QY	3.063 GASAGTAAACATTATTATCTTT 3088	1867 GTGACTGTGGCTGCAATCTTACAAACGGCTCGGTTGGGGAGAATGCGG 1926
Db	QY	7.21 GAGAGTAACATTTATCTTT 745	181 GTGACTGTGGCTGCAATCTTACAAACGGCTCGGTTGGGGAGAATGCGG 240
Db	QY	1927 AAGAGGGAGTGGCTGAGGCTGAGCCCTCACCGCTTGGCCAGGGCACTCTG 1986	

RESULT	15	
LOCUS	BI416218	
DEFINITION	BI416218-501 mRNA sequence.	
ACCESSION	BI416218	
VERSION	BI416218.1	
KEYWORDS	EST, Mus musculus (house mouse)	
SOURCE	Mus musculus	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 846)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/	
TITLE	Unpublished (1999)	
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgaps-mail.nih.gov	
COMMENT	Tissue procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.	
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (L1NL)		
DNA Sequencing by: Incyte Genomics, Inc.		
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/L1NL at: <a href="http://l1nlnci.nim.nih.gov">http://l1nlnci.nim.nih.gov</a>		
Plate: LIAM11359 row: 1 column: 19		
High quality sequence start: 20		
High quality sequence stop: 793.		
Location/Qualifiers		
FEATURES		
source	1. -846	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="czech II"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:5145762"	
	/tissue_type="pooled lung tumors"	
	/lab_host="DH10B (phage-resistant)"	
	/clone_id="NCI CGAP Lu33"	
	/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5'	
Db	1987 GGGCCCTGCTATCTCTCTACTGACTGGTCCTGAGCTATGTTAGGAGACA	2046
Db	300 GGGCTCTGCTATCTCTCTACTGGTCCTGAGCTATGTTAGGAGACA	359
Db	2047 AAAATTCAATTCTAGACTACAGCTACACCATGAGCTGAGGCTTCCTGGCTGTGAGGAGCA	2106
Db	360 AAATTCAATTCTAGACTACAGCTACACCATGAGCTGAGGCTTCCTGGCTGTGAGGAGCA	419
Qy	2107 GAAAGGCCAGTCCTCTGGGTGAGGCTGAAGTCACAAGTATCATCACCACCTT	2166
Qy	480 CCTCAATGATTCACCTTCACCTTGACTATGACATGCTGCTGAGCTGAGGAGTCAGCTAACAGTATCATCACCCACCTT	479
Qy	2227 AGTACAGGACCTGTCGCCATCTGCCCTGATGTCACCCATGTCCTCCCTG	2286
Db	540 ACTACAGCACCTGTCGCCCATCTGCCCTGATGTCACCCATGTCCTCCCTG	599
Qy	2287 GC-AAGGCCATCTGGTCACAGCTGCTGAGGAGGT-ACTGGAGGCT	2344
Db	600 GAAAGGCCATCTGGTCACAGCTGCTGAGGAGGTACCGGAGGT	659
Qy	2345 GATCTCTGAGGAGGTGAGATCGTGTGTCATC-AACCGACACCTGAGAACCTCAGTC	2403
Db	660 GATCTCTGAGGAGGTGAGATCGTGTGTCATCAGAACAGACGACTGTCAGGACCTCATGC	719
Qy	2404 CGAGAGATCACCCAGAGATGATGTCGTTGGTTCTCAGTGGGGTGTGACTCT	2463
Db	720 CGAG-AGATACCCAGAGATGATGTCGTTGGTTCTCAGTGGGGTGTGACTCT	778
Qy	2464 GCGAGGGGATCTGGTGCCCTGTGAAAGCCGGAGAAAGATGGGGATGTTCCAGG	2523
Db	779 G-CAGGGTAGCTCTGGTGCCCCCTGTGAAAGCCGGAGAAAGATGGGGATGTTCCAGG	837